مجمع اللغة العربية الأردني

الحروف العربية والحاسوب

أ.د. محمد زكي محمد خضر
الجامعة الأردنية

1996/6/22

1– مقدمة

لم يكن تطور اللغات الحية منها والمندثرة إلا سماعا، وقلما كان هناك تخطيط لتطور اللغات إلا في العصر الحديث. فاللغة العربية تحدث بها العرب وتناقلتها أجيالهم جيلا بعد جيل فأضافوا لها وأهملوا منها.

وحيث لم يكن هناك في تلك الأزمان وسائل لتسجيل الصوت ونقلها إلى الأجيال اللاحقة فإن الكتابة كانت هي الوسيلة الوحيدة (عدا المشافهة) لخزن اللغة ونقلها بدقة إلى الأجيال اللاحقة التي لم تلتق بالأجيال السالفة. ولكن عند بدء تدوين اللغة بدأت بالثبات خاصة بعد ظهور الإسلام ونزول القرآن الكريم وتدوينه، وبذلك ثبتت أسس اللغة العربية من نحو وصرف. وعلى هذا فإن اللغة العربية أُخذت أولا نطقا ثم تلتها الكتابة لاحقا، فالكتابة كان عليها أن تكون دقيقة في تمثيل النطق فأضيف لها التنقيط في القرن الهجري الأول ثم تلا ذلك التشكيل، وبذلك أمكن ضبط النص المكتوب لكي ينطق من قبل الأجيال اللاحقة وفق النطق نفسه الذي نطق به الأولون وكان الفضل في ذلك إلى نمط الكتابة العربية. فالمحافظة على نمط الكتابة يصل آخر الأمة بأولها. وهذا ما لمسه الخطاطون العرب على مر العصور حيث أضافوا للخط العربي جمالا وتفننوا في الإبداع في ذلك فظهرت الخطوط العربية المختلفة ثم أضيفت بعض الرموز كرموز الوقف والتجويد للقرآن الكريم ثم أضيفت رموز أخرى في العصر الحديث مقتبسة من اللغات الأخرى. وكل ذلك لم يغير من نمط الكتابة العربية ومواصفاتها الأصلية.

وفي خلال نصف القرن الأخير وبعد فترة وجيزة من انتشار الحواسيب أمكن إدخال اللغات التي تستخدم الحروف اللاتينية إلى الحاسوب أولا. ثم ظهرت فكرة إدخال اللغات الأخرى بعد إجراء تطويرات متعددة في تركيب وبرمجة الحاسوب. وقد حشرت الحروف العربية وما يتكون منها من كلمات وجمل حشرا ضمن البرامج لكي تكتب كما تدخل... ثم تطور ذلك...

إن الاستعمال الرئيس للحاسوب بالنسبة لمختلف اللغات لا يزال مجال الكتابة، أما مجال النطق والأصوات فقد حدثت فيه تطورات هامة خلال العقود الأخيرة إلا أنه لا يزال محدودا بالمقارنة مع مجال الكتابة. وعلى هذا فإنه سيتم التركيز على مجال الكتابة هنا على أمل أن تكون هناك فرص أخرى في المستقبل إن شاء الله تعالى لتناول موضوع النطق للغة العربية.

إن خدمة الكتابة العربية من قبل الحاسوب تستدعي جملة فرضيات أساس، من الضروري أخذها بنظر الاعتبار.

إن عمل الحاسوب يجب أن يكون وسيلة لخدمة اللغة العربية وحاكيا بدقة مختلف جوانبها وليس قاسرا لها بحيث يطوع اللغة العربية لما يناسب تصميم الحاسوب نفسه. وهذا لا يعني الجمود على الوضع الحالي للغة العربية بل يترك ذلك لعلماء اللغة وباحثيها وخطاطيها، وهؤلاء ليسوا مجبرين على أن ينحوا نحوا قسريا معينا وفق آراء واجتهادات المتخصصين بالحاسوب أو الجهات المصممة له.

أما الأمر الثاني الواجب توفره فهو أمانة تمثيل اللغة العربية من قبل الحاسوب. ويستوجب ذلك الاستيعاب الكامل لإمكانية إدخال اللغة العربية إلى الحاسوب، وإمكانية معالجتها داخل الحاسوب بتفصيل ودقة ولكل الاحتمالات المتوقعة في المستقبل، إضافة لإمكانية طباعتها بشكل اعتيادي مع الأخذ بنظر الاعتبار الطابع الجمالي المخرج من الحواسيب.

وإذا ما أخذنا بنظر الاعتبار أن الحاسوب تكون وتطور أصلا في بيئة غير عربية فإن خدمة اللغة العربية لابد وأن تكون –في الوقت الحاضر على الأقل– محدودة بالإمكانيات التقنية الحالية دون أن نهمل وضع القواعد والأسس الواجب أخذها بنظر الاعتبار لتطورات الحاسوب المستقبلية لكي يكون أكثر مطاوعة لخدمة اللغة العربية خاصة إذا ما تطورت حواسيب تحوي ذكاء اصطناعيا لا بأس به لخدمة اللغات المختلفة، فعند ذلك يصبح بالإمكان ظهور حواسيب أكثر خدمة للغة العربية مما هو متوفر الآن. وحتى يأتي ذلك الوقت فإن هناك حاجة لبحوث كثيرة في اللغة العربية وفي علم الحاسوب حول اللغة العربية خدمة للناطقين بها من أبناء هذا الجيل والأجيال القادمة.

2- نبذة تاريخية:

بعد فترة وجيزة من دخول الحاسوب إلى الحياة العملية في مختلف القطاعات وجد أن بالإمكان استخدام الطابعات السطرية Line Printers بلغات لا تستعمل الحروف اللاتينية. وعند ذلك صنعت أحزمة Belts من حروف طباعية عربية وكان ذلك في نهاية الستينات من هذا القرن. وقد احتوت تلك الأحزمة على طواقم من حروف عربية أغلبها ذات شكل واحد للحرف الواحد (عند استخدام الحرف الأول في أول الكلمة أو وسطها أو آخرها) نظرا لقلة عدد الأماكن المتوفرة على تلك الأحزمة وضرورة احتوائها على الأحرف الإنكليزية (الكبيرة والصغيرة) بجانب الأحرف العربية. وقد كانت قراءة ما يطبع من تلك الطابعات صعبة. وقد استوجبت تلك المحاولة اصطلاح مواقع رموز الأحرف ضمن جداول الرموز الحاسوبية المعروفة بـــ (ASCII Code) والتي وضعت أصلا للغة الإنكليزية من معهد التقييس في الولايات المتحدة الأمريكية. كما استوجب وضع رموز الحروف العربية منفصلة عن بعضها تفصل بينها فواصل صغيرة قد تكون مرئية للقارئ وقد تكون غير مرئية بحيث تظهر وكأنها كتابة متصلة.

وقد اتخذت الشركات المصنعة لتلك الطابعات لنفسها قرارات حددت بموجبها عدد ونوع الحروف التي تحويها الطابعات التي تصنعها. ولكنها بدافع تسويق منتجاتها، ولكونها شركات غربية لا خبرة لها باللغة العربية، بدأت بالاتصال بالمستخدمين العرب (وخاصة الجامعات ومراكز الحواسيب الكبيرة في العالم العربي) لاستمزاج رأيها في تحسين شكل الحروف الطباعية وعددها وأسلوب توزيعها بين الأحرف اللاتينية وبدأت تجري تحسينات عليها.

وبعد انتشار استخدام الشاشات المرئية قامت تلك الشركات بإنتاج شاشات للمستخدمين العرب تظهر الحروف العربية بشكل مجموعة من النقاط المضيئة. وقد حدثت نقلة نوعية مهمة في منتصف السبعينات حينما قامت إحدى الشركات بإنتاج شاشة عربية بإمكانها إظهار الحرف العربي المتكون من مصفوفة مضيئة بحيث يختلف شكله حسب موقع الحرف من الكلمة العربية ذاتيا. ثم انتقلت الفكرة نفسها بعد ذلك إلى الطابعات عند انتشار الطابعات النقطية Dot Matrix Printers. وقد كان عدد النقاط المضيئة للشاشة أو دبابيس الطابعات محدودا أولا فكان شكل الحروف رديئا. إلا أنه بعد زيادة دقة الشاشات والطابعات أصبح بالإمكان طباعة أشكال جميلة من الحروف العربية بمختلف مواقعها إضافة إلى تغيير عرض الحرف وإضافة التشكيل. وكان في خلال هذه العملية أن أضيف إلى الحواسيب ذكاء محدود لكي يتم وفقه تحديد شكل الحرف (من بين الأشكال الأربعة: الأولية والوسطية والآخرية والمنفصلة) ذاتيا حسب موقعها من الكلمة.

وقد زاد خلال هذه الفترة الضغط على الجهات العربية المسؤولة عن المقاييس العربية وكثرت المناقشات والمداولات في الندوات والمؤتمرات العربية حول الحواسيب.

3- المواصفات القياسية العربية:

ساهمت عدة منظمات ومؤسسات عربية ودولية في بلورة واختيار مواصفة قياسية عربية مناسبة للحاسوب. ومن تلك المنظمات[1] : المنظمة العربية للتربية والثقافة والعلوم، ومعهد اللسانيات في المغرب، والمنظمة العربية للمواصفات والمقاييس، والمنظمة السعودية للمواصفات والمقاييس، والمركز القومي للحاسبات الإلكترونية في العراق، والاتحاد العربي للاتصالات، ومعهد الكويت للأبحاث، وشركة أليس بكندا، ومكتب ما بين الحكومات للمعلوماتية، والاتحاد الأوربي لمصنعي الحاسوب.

وفي عام 1981 شكلت المنظمة العربية للمواصفات والمقاييس التي كانت تابعة للجامعة العربية ومقرها في عمان، لجنة لتحديد مواصفة قياسية عربية للحروف العربية في حقل المعلومات. وقد أصدرت المنظمة المواصفة القياسية رقم 449 (ASMO 449) وسجلت عالميا بـــ (ISO/9036) وقد احتوت 120 حرفا للاستعمال في حقل معالجة المعلومات وقد أُقرت في كانون الثاني عام 1985. ويبين الشكل (1) توزيع الحروف على المواقع وهي تحوي على سبعة أرقام ثنائية (7 bits).[2][3]

وفي تشرين الأول عام 1985 أقرت المنظمة مواصفة قياسية أخرى برقم 662 (ASMO 662) تتطابق مع المواصفة الأولى في مواقع الأحرف مضافا إليها حروف التحكم وكانت ذات ثمانية أرقام ثنائية .(8 bits)

وبعد عام من ذلك تقريبا (تشرين الثاني عام 1986) أقرت المنظمة المواصفات القياسية 708 (ASMO 708) لتبادل المعلومات على الحاسوب بثمانية أرقام ثنائية وسجلت عالميا تحت رقم .(ISO/8895-6)

ويبين الشكل (2) توزيع الأحرف العربية على النصف العلوي من الجدول فيما أبقيت الأحرف الإنجليزية على النصف السفلي من الجدول كما هي. ويلاحظ في هذا المواصفة أن الحروف العربية متطابقة في مواقعها مع المواصفة 449 وكذلك أدوات التشكيل الثمانية وهي على التعاقب:تنوين الفتح ثم تنوين الضم ثم تنوين الكسر ثم الفتحة فالضمة فالكسرة فالشدة فالسكون.

أما الأرقام العربية فقد احتلت في المواصفة 708 مواقع الأرقام الإنجليزية نفسها وليس من موضوعنا بحث مواقع أحرف التحكم في هذه المواصفات.

ونظراً لتأخر صدور هاتين المواصفتين القياسيتين أكثر من عشر سنين على بدء استخدام اللغة العربية في الحاسوب فقد وضعت الشركات المنتجة للحواسيب مواصفاتها الخاصة بها مما جعل من الصعب عليها التراجع عنها بسبب ما قد يكلفها ذلك من تكاليف عالية. لذلك استمر وجود أطقم عديدة من أنظمة الحروف بحسب الشركات المنتجة للحواسيب.

4- أطقم الحروف العربية

يبين الملحق (1) معظم الأطقم للحروف العربية الشائعة. وفيما يأتي ملاحظات عن أكثر هذه الطواقم انتشاراً:

4-1 المواصفتين القياسيتين للمنظمة العربية للمواصفات والمقاييس المرقمتين 449 و 708

وقد سبقت الإشارة لهما. وتمتازان بما يأتي:

أ) وجود موقع لكل حرف من الحروف الثماني والعشرين.

ب) وجود ستة مواقع للهمزة بأشكالها المختلفة ء أ إ ؤ ئـ آ

ج) وجود موقع للتاء المربوطة وآخر للألف المقصورة.

د) وجود ثمانية مواقع للتشكيل هي ً ٌ ٍ َ ُ ِ ّ ْ وحسب الترتيب المبين.

هـ) وجود موقع للكشيدة (التطويلة) بين الحروف المتصلة.

4-2 مجموعات أطقم الحروف الخاصة بالدوس العربي:

أصدرت شركة أليس تعريباً لنظام التشغيل وعرف بالدوز العربي، وقد استندت الشركة إلى أطقم المواصفة القياسية العربية 708 ثم أجريت تعديلات على المواصفة القياسية 449 وأعطيت الرقم 709 وبعد ذلك أنتجت طواقم أخرى برقم 710 و 720.

فالطاقم المرقم 709 يتطابق مع المواصفة 449 مضافا إليها مجموعة من التشكيل المزدوج على التوالي : ًّ ٌّ ٍّ َّ ُّ ِّ .

أما الطاقم 710 فلا يحوي تشكيلاً وتنقسم الحروف العربية فيه إلى سلسلتين الأولى تحوي أشكال الهمزة والألف والباء والتاء المربوطة، يليها سلسلة من الرموز غير الحرفية ثم تأتي السلسلة الثانية من الحروف العربية من التاء إلى الياء والكشيدة ولا يحوي هذا الطاقم على التشكيل.

أما الطاقم 720 فقد احتوى على التشكيل الأساسي بثمانية مواقع مع الحروف العربية موزعة على عدة سلاسل، تتوزع بينها رموز أخرى غير الحروف العربية.

4-3 مجموعات النافذة العربية

أصدرت شركة 01 البحرينية نظاما للتعريب يعتمد على أساس شفافية الحروف العربية بحيث يتواءم مع أطقم الحروف اللاتينية وقد أصدرت الشركة عدة طواقم للحروف أشهرها المرقم 711 ويحوي الحروف العربية إضافة إلى التشكيل وحسب الترتيب الآتي:

أ) التشكيل ْ ّ َ ً ُ ٌ ِ ٍ .

ب) الهمزات بأشكالها الستة ثم الألف والباء والتاء المربوطة.

جـ) الحروف من التاء إلى الغين.

د) الكشيدة.

هـ) الفاء إلى الميم.

و) النون والهاء.

ز) الواو والألف المقصورة والياء.

و) يتخلل مواقع هذه الحروف رموز بيانية.

4-4 أطقم صخر العربية

أصدرت شركة العالمية (صخر) تعريبها الخاص بها وقد احتوى طاقم الحروف العربية لها على ما يأتي:

أ) أدوات التشكيل الأساسية الثماني ً ٌ ٍ َ ُ ِ ّ ْ .

ب) الهمزات الستة ثم الألف والباء والتاء المربوطة.

جـ) الحروف الأخرى من التاء وحتى الياء ثم الكشيدة.

ويتخلل الحروف (بين التاء المربوطة والتاء) الرموز البيانية.

4-5 طاقم شركة ا.ب.م المرقم 864

أصدرت شركة ا.ب.م الأمريكية طاقما خاصا بها أسمته:

IBM Arabic National Supplement 864 ويحوي هذا الطاقم على ما يأتي:

أ) مجموعات اللام ألف وهي ثمانية مواقع لأ ـلأ لإ ـلإ لا ـلا لآ ـلآ.

ب) التعريقة .

جـ) الهمزات على الألف آ ـآ أ ـأ إ ـإ.

د) الشدة ّ ـّ.

هـ) الحروف العربية بشكلين لكل حرف بحيث يمكن كتابة أشكال الحروف مع التعريقة.

وقد تم توزيع الحروف والرموز البيانية متداخلة بينها على غير تسلسل معين وبتجميع بعضها مع بعض أحيانا.

6-4 طاقم شركة مايكروسوفت

بعد انتشار نظام النوافذ windows المنتج من شركة مايكروسوفت الأمريكية أصدرت الشركة تعريبا له وقامت باستحداث طاقم خاص بها ويحوي هذا الطاقم على:

ا) الأشكال الستة للهمزة.

ب) الحروف العربية من الألف إلى الضاد تتخللها التاء المربوطة بين الباء والتاء.

جـ) حروف الطاء والظاء والعين والغين تليها الكشيدة ثم الفاء والقاف والكاف.

د) حرف اللام.

هـ) حروف الميم والنون والهاء والواو.

و) حرفي الألف المقصورة والياء.

ز) التشكيل ً ٌ ٍ َ ُ ِ ّ ْ تتخللها رموز أخرى.

وقد بدأ هذا الطاقم بالانتشار أكثر من غيره بعد طغيان استعمال النوافذ هذا ويخشى أن يبقى هو السائد، طاغيا وملغيا لجميع الأشكال الأخرى أو معظمها.

تبين الأشكال 3-9 توزيع الحروف العربية لعدد من الأطقم المذكورة أعلاه(6).

وتجدر الإشارة إلى أن هناك إمكانية في بعض الأنظمة (أو البرامج الجاهزة) التي يمكن بواسطتها التحويل من طاقم لآخر كما هو الحال في نظام ويندوز الخاص بمايكروسوفت ونظام النافذة.ولكن بالطبع فإن عملية التحويل هذه لا تخلو من مشاكل نظرا لعدم تكافؤ الطواقم تماما في محتوياتها.

5- أنظمة تعريب الحاسوب

نظرا لما للحروف العربية من علاقة بأنظمة تعريب الحاسوب التي ظهرت لحد الآن فسنعطي فكرة مبسطة عن أنظمة التعريب التي عرفت لحد الآن، فإن من يستعمل اللغة العربية يجابه عددا من المعضلات التي أخذت أنظمة التعريب على نفسها مهمة حلها. ومن هذه المعضلات:

أ) الكتابة من اليمين إلى اليسار بخلاف اللغات التي تستعمل الحروف اللاتينية التي تكتب من اليسار إلى اليمين.

ب) إمكانية إدخال نص عربي ضمن الكتابة الإنكليزية (أو الفرنسية)، وبالعكس إمكانية إدخال نص إنكليزي (أو فرنسي) ضمن النص العربي.

جـ) الاختيار بين الأرقام العربية الغبارية المستعملة في المشرق العربي والأرقام المستعملة في المغرب االعربي وأوربا.

د) الشفافية، والمقصود بها المحافظة عند التعريب على إمكانية استخدام البرامج الإنكليزية الأصلية بحيث لا تؤثر الحروف العربية سوى على مواقع بعض الرموز البيانية.

هـ) التعامل مع الحروف الحيادية (بين العربية واللاتينية) مثل الفراغ والعلامات والأرقام (كقيم وليس كشكل) بشكل منطقي ومقبول.

وقد قامت الجمعية العلمية الملكية بعمان عام 1989 بإجراء دراسة مقارنة عن أنظمة التعريب(7) بين أحد عشر نظام تعريب شائع، بينها: تعريب ا.ب.م، والمساعد العربي، والأمير، والدوز العربي، والنافذة، والأستاذ، وغيرها. وقد ألقيت الدراسة كبحث في المؤتمر الأول للحاسوب في الكويت الذي عقد بين 27-89/3/29. ويجدر بالذكر نظامان آخران مهمان للتعريب هما النظامان الذين أنتجتهما شركة صخر ونظام النوافذ windows الذي أنتجته شركة مايكروسوفت.

وقد طورت الأولى نظامها الذي كان أصلا يعتمد على أجهزة صخر فقط ثم أنتجت النوافذ العربية مستندة إلى شفافية خاصة بها. أما نظام النوافذ الذي أنتجته شركة مايكروسوفت فقد أصدرت منه نظاما معربا ثم أنتجت طبعة جديدة عام

1995 سميت النوافذ 95 ويتوقع أن تصدر تعريبا لها في وقت لاحق من هذا العام. أما شركة أبل ماكنتوش فقد أصدرت تعريبا خاصا بها على أجهزتها غير المتوافقة مع ا.ب.م واستندت في طاقم الحروف العربية إلى توزيع خاص بها لكن يقرب من الطاقم المرقم 711. وتمتاز برامج هذه الشركة بقابلية إظهار الحروف العربية بدقة عالية مما مكن من إنتاج خطوط عربية جميلة بواسطة أجهزة هذه الشركة.

ويلاحظ بين أنظمة التعريب أن مسألة ترتيب الحروف الهجائية يعتمد على الطاقم المستعمل حيث أن كل هذه الطواقم قد اعتبرت التشكيل عبارة عن حروف، لذلك فإن ترتيب الحروف عند الكتابة المشكولة يعتبر التشكيل حرفا يؤخذ بعين الاعتبار عند الترتيب. ومن ذلك ينتج أن الطاقم نفسه يحدد أسلوب الترتيب. وإذا ما كانت الحروف مبعثرة في الجدول فإن الترتيب سيكون ذي إشكالات عديدة وخاصة إذا ما أخذ شكل الحرف حسب موقعه في الجدول كما هو الحال في الطاقم الخاص بشركة ا.ب.م المرقم 864. كما أن مواقع الحروف الخاصة (الهمزة والتاء المربوطة والألف المقصورة والكشيدة) كلها تختلف من طاقم لآخر مما يولد ترتيبا مختلفا عند استعمال طاقم معين عن الطاقم الآخر.

ويجدر بالإشارة إلى أنه لحد الآن لم يتم عمل برنامج للتحاور باللغة العربية مع الشبكة العالمية Internet وهذا البرنامج له ضرورة بالوقت الحاضر لتطوير التحاور ونقل البريد الإلكتروني بدون استخدام أنظمة التعريب الأخرى.

6– الرمز الدولي الموحد Unicode

إن أكثر نظام انتشارا في ترميز الحروف والرموز هو نظام ANSI للحروف اللاتينية (وهو موضوع من قبل معهد المقاييس الأمريكية

American National Standards Institute) وهو مناسب للغات المستخدمة للحروف اللاتينية. وقد تمت استعارة قسم من المواقع المخصصة للزخرفة Graphics لوضع الحروف العربية (أو غيرها من الحروف عند استخدام أنظمة ثنائية اللغة) مع المحافظة على مواقع الحروف اللاتينية.

وقد وجد أن هناك لغات كالصينية والكورية واليابانية يزيد عدد الحروف فيها عن عشرة آلاف حرف وبذلك لا يصلح نظام ANSI لها. كما أن أي نظام ثنائي اللغة ليس بينهما إحدى اللغات المستخدمة للحروف اللاتينية يعني أن تحل حروف إحداهما محل الحروف اللاتينية وبذلك تغير موقعها المستخدم عند استخدامها مع الحروف اللاتينية. هذا بالإضافة إلى أن هناك كثيرا من الرموز الرياضية والتقنية الحديثة التي وجدت ضرورة لتحديد مواقع لها في طواقم الحروف. وكل ذلك يزيد عن العدد 256 المخصص ضمن بايت byte واحد ذي ثماني أرقام ثنائية (8 bits).

ولغرض تلافي هذا الإشكال فقدت اتفقت عدة شركات عالمية على تشكيل منظمة عالمية غير ربحية سميت منظمة الرمز الموحد Unicode لغرض تعريف نظام قياسي عالمي يمكنه أن يضم كافة الحروف المستخدمة في كافة لغات العالم الحية (وحتى المنقرضة منها كالمصرية القديمة والسنسكريتية).

كما أن المنظمة العالمية للتقييس ISO بدأت بتطوير نظام للتقيس المناسب لذلك. وهكذا تكون الرمز العالمي الموحد Unicode.

6–1 اللغات غير العربية المستخدمة للحروف العربية

يبين الشكل (10) البلاد المستخدمة للحروف العربية وتشمل بالإضافة للبلدان العربية أفغانستان، وإيران، ومالطا، والهند، وباكستان، وإقليم كينيانج في الصين، وكازخستان، وتترستان، وهاتان اللغتان بالرغم من صلتهما الوثيقة باللغة التركية إلا أنهما لا تزالان تستخدمان الحروف العربية رغم أن تركيا تستخدم الحروف اللاتينية.

أما ماليزيا، وإندونيسيا، والصومال فإنها رغم استخدامها للحروف اللاتينية إلا أن هناك صحفا تستخدم اللغة العربية في ماليزيا وجاوا. وفي إفريقيا فإن اللغات البربرية، والهوسا، والسواحلية تستخدم الحروف العربية أحيانا.

ويبين الشكل (11) أبجديات بعض اللغات غير العربية التي تستخدم الحروف العربية.

فبالرغم من أن الأبجدية العربية تحوي 28 حرفا، إلا أن اللغات التي تستخدم الحروف العربية تحوي أبجدياتها حروفا فوق الحروف العربية الثمانية والعشرين وبعضها مشتق منها. فالفارسية تحوي 32 حرفا، والأُردية تحوي 36 حرفا، والسندية تحوي 52 حرفا. كما تحوي لغة الباشتو المستخدمة في أفغانستان وإقليم بلوجستان في الباكستان على 10 حروف إضافية، كما تستخدم لغة اليوغر المعروفة في شمال غرب الصين على 10 حروف إضافية .... وهكذا.

6-2 الحروف العربية وما يشتق منها في الرمز العالمي الموحد

أصدرت منظمة Unicode الرمز العالمي الموحد الذي يحوي على 65536 حرفا وقد تم تخصيص حوالي 34000 حرفا منها للغات الحية. ويبين الشكلان (12) و (13) جدولي الحروف العربية الأصلية والحروف المشتقة منها المستعملة في غير العربية. ويلاحظ أن الجدول يحوي ستة أشكال للهمزة والحروف الأبجدية الثماني والعشرين يتخللها حرف التاء المربوطة بين الباء والتاء وحرف الألف المقصورة بين الواو والياء. والكشيدة قبل الفاء. وأدوات التشكيل الثمانية ً ٌ ٍ َ ُ ِ ّ ْ . كما تحوي على الأرقام المستعملة في المشرق العربي وتحوي كذلك حرفين آخرين لهما أهميتهما وهما الألف الخَنجرية وهمزة الوصل. أما باقي الأحرف في الشكل (12) فهي ليست مستخدمة باللغة العربية بل باللغات الأخرى المستخدمة للغة العربية.

ويتضح أن الأحرف هذه تحتل المواقع من 0600 في الجدول العالمي الموحد إلى F 067 للجدول الأول ومن 0680 إلى 06FF للجدول الثاني كما خصصت المواقع من 0700 لغاية 08FFلأية تطورات في الحروف العربية وذات الصلة بها وهي الآن فارغة (احتياطية).

7- الوضع الراهن للحروف العربية في الحاسوب

هناك أربعة مجالات مختلفة للعلاقة بين الحروف العربية والحاسوب. وهي مجال إدخال المعلومات ومجال إخراجها ومجال معالجتها والجانب الجمالي للحرف العربي. ورغم الترابط بين هذه المجالات إلا أن لكل منها خصوصيات خاصة بها.

7-1 إدخال المعلومات والحرف العربي

هناك طريقتان لإدخال المعلومات المكتوبة للحاسوب هما عن طريق لوحة المفاتيح keyboard وتمييز الحروف ضوئيا المعروفة بـــ OCR. ففيما يتعلق بلوحة المفاتيح يجب أن تحوي هذه اللوحة على كل الإمكانيات الخاصة بالإدخال فهي يجب أن تحوي الحروف الثماني والعشرين والهمزة بأشكالها والتاء المربوطة والألف المقصورة وأدوات التشكيل الثمانية والأرقام العربية وربما الألف الخنجرية وهمزة الوصل. وليس هناك أية لوحة مفاتيح تحوي هذين الحرفين في الوقت الحاضر. إلا أنه يلاحظ أن هناك إمكانية في لوحات مفاتيح متعددة لإدخال اللام ألف جملة واحدة بل وبأشكال متعددة: لا لأ لإ لآ. ورغم أن برامج التعريب تأخذ بنظر الاعتبار تحويل اللام ثم الألف إلى لام ألف غالبا، إلا أن إدخال اللام ألف جملة واحدة عملية لا بأس بها في تسهيل الإدخال. ولا بأس أن تحوي لوحات المفاتيح أية مجموعات من الحروف لتسهيل الإدخال مثل ألف لام التعريف.

ويلاحظ في الإدخال أنه يجب أن يكون متوافقا مع الرموز التي يتعامل بها الحاسوب داخليا إلا أنه لا يشترط شكل واحد للوحة المفاتيح. وقد لوحظ هناك اختلافات في مواقع بعض الحروف كالدال والذال والطاء وأدوات التشكيل بين لوحة مفاتيح وأخرى. ولذلك ينبغي أن يكون هناك توافق أكبر في لوحات المفاتيح الحاسوبية تسهيلا لعملية الإدخال. ولكن الأكثر ضرورة هو تحديد الحد الأدنى من الحروف المتفق عليها ومواقعها.

أما أجهزة تمييز الحروف العربية OCR فقد ظهرت مؤخرا أربعة برامج منها[10]. وهذه الأجهزة تقوم بعملية مسح (scanning) للصفحة المطلوب تمييز حروفها ثم تقوم بمطابقة الحروف حرفا حرفا مع ما مخزون بداخلها من نماذج سبق أن دربت عليها. وبعضها يحوي إمكانيات أوسع من ذلك دون الحاجة إلى تدريب.

لقد تطورت أجهزة تمييز الحروف اللاتينية إلى درجة إمكانية تمييز الكتابة اليدوية. أما في اللغة العربية فان البرامج المتوفرة لحد الآن تنحصر دقتها في الخطوط الطباعية الجيدة والحديثة الطباعة. أما الكتابة القديمة فإن تمييزها لا يزال صعبا.وليس هناك أية برامج بإمكانها تمييز الحروف المكتوبة بخط اليد بعد.

يمكن النظر إلى الحروف العربية كمجموعات تتألف من 15 مجموعة كما مبين في الشكل (14) فالجيم والحاء والخاء مثلا لا تمييز بينها إلا بالنقط. وهذا التقسيم يسهل عملية التمييز، وبالطبع اذا ما أريد له ان يكتمل فيجب أن يحوي الهمزات وأشكال اللام ألف والتشكيل وغير ذلك.

إن تمييز الحروف العربية ذو علاقة وثيقة بأشكال طباعة الحروف مطبعيا وهذا يخضع لمسألة جمال الخط العربي والتي ستناقش فيما بعد.

7-2 إخراج الخط العربي

إن متطلبات إخراج الخط العربي من حيث عدد الحروف لا يختلف عن متطلبات الادخال فالقراءة هي لخط مكتوب والكتابة هي عكس عملية القراءة.

أما من ناحية جمال الخط العربي فهي مسألة أخرى.

إلا أن هناك جملة أمور تتعلق بأسلوب الكتابة العربية. فقد كانت الحروف المطبوعة أول استخدام الحاسوب ذات شكل واحد مع إضافة التعريقات أحيانا، ثم أدخل عليها تحسين جوهري بتمييز شكل الحرف في أول الكلمة أو وسطها أو آخرها أو الحرف المنفصل. وكان عرض الحرف مبدئيا متساو فالصاد المنفصلة كانت تحتل عرضا مساويا للباء الوسطية مثلا. وبتطور تقنية الطباعة (والإخراج على الشاشة) أمكن تخصيص عرض مختلف حسب الحاجة لكل شكل من أشكال الحروف.

أما التشكيل فقد كان يعتبر حرفا له عرض الحروف الأخرى سواء كان مع الكشيدة أو بدونها (مثلا الفتحة وسط الكلمة تحتاج إلى كشيدة أما الفتحة على حرف غير متصل من جهة اليسار مثل الواو فلا تحتاج إلى كشيدة).

وهكذا برزت الحاجة إلى صورتين للتشكيل إحداهن مع الكشيدة والأخرى بدونها. إلا أن تطور تقنية معالجة المعلومات بتغيير عرض الحرف مكن من وضع التشكيل فوق الحرف السابق مباشرة، بل وأمكن دمج الشدة مع الفتحة والضمة والكسرة ذاتيا، فكلما وردت شدة بعدها فتحة مثلا دمجتا ووضعتا فوق الحرف السابق.

وهكذا يظهر باب الإخراج مختلف عن الإدخال. ففي الإدخال يجب إدخال الحرف ثم الشدة ثم الفتحة بينما عند الإخراج يظهر الحرف وفوقه الشدة وفوقها الفتحة. وهذا تطور لا بأس به في جمال الكتابة العربية.

7-3 جمال الكتابة العربية

مع تطور تقنية الطباعة وإظهار الكتابة على الشاشة بدقة أعلى، ظهرت تقنيات خاصة بالحروف (اللاتينية وغيرها). وأحد هذه التقنيات ما سمي بالنوع الحقيقي للفونط True Type Font ويرمز له بـــ (TTF) وتستند هذه التقنية على أساس خزن المواصفات العامة للحرف المطلوب إظهاره مهما كان حجم الطباعة أو الإظهار على الشاشة المطلوب فالشكل يبقى كما هو مهما كان الحجم. ويمكن بواسطة هذه التقنية إظهار حروف متلاصقة وبأية أشكال يرتأيها الخطاطون. يبين الشكل (15) أحد جداول الخطوط العربية لشركة مايكروسوفت المسمى بالخط العربي التقليدي، ويلاحظ فيه وجود موقع لمجموعات من الحروف المتلاصقة المكونة من حرفين مثل (محـــ، ـــبن، نجـــ.. الخ). وهكذا بإمكان الخطاط أن يضع ما يريد من مجموعات حرفية تزيد من جمال الخط، ولكن يجب أن تبرمج بنسق معين لكي يقوم الحاسوب بإخراج هذه المجموعة من الحروف بالصيغة المطلوبة كلما تمت مصادفتها أثناء الكتابة.

وقد ظهرت مؤخرا أشكال جميلة للكتابة العربية بخطوط مختلفة قامت باصدارها بيوت البرمجة العربية في لبنان والسعودية ومصر وأخرى في الدول الغربية.

يبين الشكل (16) سورة الفاتحة مكتوبة بالرسم العثماني للقرآن الكريم، ويلاحظ فيها وجود أكثر من نوع للميم في آخر الكلمة وكذلك النون. إضافة إلى أي شكل من أشكال الحروف كالألف الخنجرية وهمزة الوصل بل وإضافة علامات الوقف والتجويد.. الخ. كما أن الهمزة المكسورة توضع تحت الكلمة (مثل الملـــبكة). إلى غير ذلك من خصوصيات كتابة المصحف.

7–4 معالجة اللغة العربية داخليا في الحاسوب

سبق وأن أوضحنا أن الحروف العربية يرمز لها برمز واحد داخل الحاسوب وهذا الرمز يتطابق مع موقع واحد على لوحة المفاتيح. أما الإخراج فيمكن أن يظهر مختلفا حسب موقع الحرف من الكلمة. ومن ذلك يظهر أنه ليس من الضروري تطابق شكل الحرف عند الإدخال مع ما مخزون داخل الحاسبة مع ما يطبع.

فاللام ألف ليس لها موقع في جدول الحروف العربية ولكنها تظهر كحرف واحد عند الطباعة وقد يكون إدخالها دفعة واحدة أو بالكبس على اللام ثم الألف حسب لوحة المفاتيح. وهذا المقدار من الذكاء الداخلي موجود في بعض أنظمة التعريب حاليا.

إلا أن هناك امور كثيرة لا تزال بحاجة إلى عناية حيث أن هناك قصورا في معالجتها وسنأتي على ذكر بعض منها.

7–4–1 ترتيب الحروف العربية

عمل بعض علماء اللغة العرب القدامى على ترتيب حروف المعاجم العربية حسب مخارج الحروف. وهناك لا شك الترتيب الهجائي المعروف (ا ب ت ...) والأبجدي (ا ب ج د هــ و ز ...).

لقد وضعت معظم طواقم الحروف العربية التي سبق ذكرها بالاستناد إلى الترتيب الهجائي (ا ب ت ...) إلا أن هناك اختلافات متعددة فيما يتعلق بوضع فجوات بينها تضم رموزا أو حروفا غير عربية كما أن الحروف الإضافية (كالتاء المربوطة والألف المقصورة وأشكال الهمزة .. الخ) كلها غير مجمع على عددها ولا على مواقعها في الطواقم. كما أن التشكيل الذي تعتبره الحواسيب حروفا مستقلة هو الآخر موضوع بتسلسل مختلف بين طاقم وآخر.

إن تأثير وضع الحروف حسب تسلسل معين يؤثر في ترتيب كلمات تضم تلك الحروف. فإن موقع التاء المربوطة (قبل التاء الطويلة أو بعدها) يؤثر على ترتيب كلمات مثل (قائمة، قائمتان).

كما أن ترتيب الهمزة يؤثر في كلمات مثل (فاني، فأنك، فإنما، فآت، فئته) فرغم أن الكلمات هذه تحوي ألفا أو همزة تلي الفاء إلا أن ترتيبها يخضع لتسلسل ترتيب أنواع الهمزات في طواقم الحروف المختلفة. وكذلك الحال بالنسبة للتشكيل، فالفتحة مثلا تسبق الضمة في بعض الطواقم، ويعني ذلك أن الكلمات المحتوية على حرف مفتوح يسبق الكلمة نفسها التي فيها حرف مضموم حتى ولو كان الحرف الذي يلي المضموم قبل الحرف الذي يلي الحرف المفتوح في الكلمتين. مثال على ذلك كلمتي (قَفَلَ تسبق قُتِل لأن القاف الأولى تليها الفتحة ثم الفاء بينما القاف الثانية تليها الضمة ثم التاء والفتحة قبل الضمة رغم أن التاء تأتي قبل الفاء!! رغم أن ترتيب هاتين الكلمتين يكون صحيحاً بدون تشكيل). والحال تتكرر بالنسبة للشدة فهي تعتبر حرفاً مستقلاً شأنه شأن أدوات التشكيل الأخرى ولا يعتبر تكراراً للحرف الذي سبقه.

لا شك أن بعض المشاكل التي صاحبت هذه الوضعية هي مستحدثة ليس لها حلول لغوية قديمة، حيث نجم بسبب التطورات التقنية الحديثة وبعضها ربما له حلول لغوية قديمة لكن عدم إحياء تلك الحلول جعل الاجتهادات المختلفة الحديثة تتجه اتجاهات متباينة.

7-4-2 أنواع الألف

في الأطقم المستعملة في الوقت الحاضر هناك الألف اللينة والألف المقصورة وفي الرمز الدولي الموحد هناك اختيار للألف الخنجرية التي لم تستعمل بعد في الأنظمة المشهورة وهي مطلوبة في الكتابة غير المشكولة لكلمات عديدة مثل (الله، الرحمن، إله، ذلك، هؤلاء، هذا، .. الخ) كما أن هناك خلطاً في الكتابة العادية بين الألف اللينة والألف المهموزة عند غياب التشكيل. إن عمل أي ترتيب أو إحصاء لأعداد الحروف في نص معين سيعطي نتائج متباينة حسب الطاقم المستعمل وعند الخلط مع الألف المهموزة. كما أن الألف المقصورة والألف اللينة الآن تعتبران حرفين مستقلين لا علاقة بينهما.

إن دخول النحو والصرف العربي في مجال معالجة اللغة العربية في الحاسوب سوف يبرز مشاكل عند قلب الحرف نفسه من شكل لآخر، وكمثال على ذلك كلمة عصى عند إضافته إلى ضمير الغائب تصبح عصاه.

كما أن الأنظمة المتوفرة في الوقت الحاضر يمكن أن تؤدي إلى لبس عند الضغط علىمفاتيح (ع ص ى ه) سوف تظهر الكلمة (عصىه) وذلك نتيجة خلو الحاسوب عن ذكاء يتعلق بالأف المقصورة واتصالها مع غيرها .

كما أن هناك نقصاً في همزة الوصل حيث تعتبر ألفاً.

أما (ألف واو) الجماعة التي لا تلفظ بل تكتب فقط، فهي تعتبر ألفاً عند الترتيب والتشكيل والإحصاء وهو أمر قد يحتاج إلى تفكير وإعادة نظر.

7-4-3 أنواع الهمزات

تتوفر في معظم الطواقم الآن الهمزة المنفردة ء والهمزة على كرسي الياء بأشكالها الأربعة ئـ ـئـ ـئ ئ والهمزة على كرسي الواو بشكليها ـؤ ؤ والهمزة مع الألف بأشكالها الأربعة أ إ ـأ ـإ فالهمزة فوق الألف تستعمل للضم والفتح، وتحت الألف للكسر. وبذلك يحصل التمييز بين الهمزات وفق تشكيلها بخلاف باقي الحروف حتى وإن كان التشكيل غائباً! وهو أمر فيه بعض الغرابة.

أما الهمزة الممدودة بشكليها آ ـآ فتعني حقيقة همزة مع ألف لينة (ء ا) كما أن الهمزة فوق وتحت الألف عند ورودها بعد اللام تصبح لام ألف كما سيرد ذكره.

7-4-4 اللام ألف

تتفق طواقم الحروف العربية تقريباً (فيما عدا ا.ب.م 864) على أن اللام ألف تعتبر في الطاقم حرفين حتى وإن أدخل الحرفان بواسطة مفتاح واحد معنون (لام ألف). وهناك عدد من المفاتيح المخصصة للام ألف هي لا، لأ، لإ، لآ وتظهر هذه الحروف بشكلين (منفصل ومتصل أي لا و ـلا و ـلأ و ـلإ الخ). إلا أن كل أنظمة التعريب الحالية تفشل في خدمة اللام الف خدمة مناسبة عند وجود تشكيل على اللام فإذا ما أدخلت الحروف (ل َ ا) فالناتج على الحواسيب حاليا يظهر (لَـــا) وهو شكل غير مقبول ويستدعي الإصلاح رغم أن إدخال (ل ا) يعطي الشكل المقبول (لا) في معظم أنظمة التعريب الحالية.

ولإصلاح الوضع ينبغي دراسة وضع التشكيل على اللام والألف وأحوال الألف المهموزة وغير المهموزة، وعند إمعان النظر في ذلك يتضح أن هناك حالات كثيرة ولعل جدول (1) يوضح كافة الحالات سواء الممكنة منها في اللغة العربية أو غير الممكنة التي أشير إليها بالرمز *.

ولغرض معالجة تشكيل اللام ألف فإن من الجدير أخذ ما يأتي بنظر الاعتبار:

أ) إبقاء المفاتيح الأربعة الخاصة باللام ألف كما هي لاستعمالها عند عدم التشكيل.

ب) عندما يراد إدخال التشكيل يدخل اللام والألف كحرفين منفصلين بينهما التشكيل المطلوب مثلا لكتابة كلمة كلا تدخل (ك َ ل ّ َ ا) فتظهر كلا أي الشدة والفتحة فوق اللام وليس الألف. وهكذا بالنسبة للهمزة فوق الألف أو تحته مع التشكيل الخاص بها.

جـ) يقترح على مصممي الأنظمة أن تدخل معالجة اللام والألف بحيث تسمح للمدخل ان يدخلها كما يشاء من ناحية التشكيل على أن يترك للمدقق الإملائي أو برامج التشكيل الآلي إلغاء الحالات غير الممكنة في اللغة العربية.

7-4-5 التاء المربوطة

التاء المربوطة هي تاء عند اتصالها بضمير بعدها (مثل نكرة نكرتين) ولكن بعد انقلاب التاء المربوطة إلى تاء طويلة لا يعرف الحاسوب أصلها. فإذا ما أجريت لها عمليات تحليل صرفي أو نحوي فان أصلها سيكون مجهولا ما لم يتم الرجوع إلى المعاجم. كما أن المعالجة الحالية غير مناسبة عند إدخال التاء المربوطة وبعدها حرف فإن الكتابة تكون غير مناسبة (مثل نكرةين).

لذلك فإن من الضروري إدخال نوع من الذكاء لمعالجة التاء المربوطة شأنها شأن الألف المقصورة.

8- المناقشة والتوصيات

8-1 تطوير الوضع الحالي

مرت رحلة إدخال الحروف العربية إلى الحاسوب بجملة تطورات ظهرت أثناءها اجتهادات شتى في أسلوب معالجة الحروف العربية، وقد ساهمت الجهات العربية في الوصول إلى بعض الحلول المناسبة إلا أن الشركات الأجنبة الكبرى لا يزال لها التأثير الكبير في فرض المنهج الذي تتبعه في معالجة الحروف العربية وفلسفة تعريب الحاسوب بشكل عام ونظرا لعدم وجود مرجعية عربية قياسية واحدة، فإنه ربما يمر زمن ليس بالقصير حتى تصل الأمور إلى أسلوب مناسب. إن عدد الحروف العربية الأدنى المطلوب لمعالجة الحروف العربية يجب أن يشمل:

أ) الحروف العربية الثماني والعشرين بأشكال كتابتها المختلفة أول الكلمة ووسطها وآخرها ومنفصلة.

ب) التاء المربوطة والألف المقصورة.

جـ) الهمزة بأشكال كتابتها المختلفة ء ئـ ـئـ ـئ ئ ؤ ـؤ أ إ آ ـأ ـإ ـآ.

د) الكشيدة.

هـ) همزة الوصل ٱ ـٱ وهي غير مستعملة لحد الآن.

و) الألف الخنجرية وهي غير مستعملة لحد الآن .

ز) أدوات التشكيل وهي ً ٌ ٍ َ ُ ِ ّ ْ مع إمكان الجمع بين الشدة وأدوات التشكيل التي سبقتها.

ح) اللام ألف ليست حرفاً واحدا له موقع، لكنه مكون من حرفين في النص. أما الإدخال فلا بأس بان يدخل بالكبس على مفتاح واحد. ويمكن أن يكون له أشكال متعددة لا ـلا لأ ـلأ لإ ـلإ لآ ـلآ. كما أن تشكيل اللام والألف يجب أن يكون بالإمكان إدخاله وإظهاره بشكل مناسب فهذا غير موجود لحد الآن.

إن المناقشة أعلاه ما هي إلا تثبيت للوضع الحالي مع إجراء التحويرات والتعديلات المناسبة والضرورية عليه، حيث أن التشكيل في كل هذه الحالات يعتبر حرفاً مستقلا بل وإن الشدة مع الفتحة بعد الحرف ستجعل الحرف المشكول ثلاثة حروف داخل الحاسوب.

2-8 مقترح لمعالجة التشكيل

كما أن الحرف العربي المشكول يعتبر في اللغة العربية حرفاً واحداً، ألا يحق لنا أن نتساءل لماذا لا يمكن اعتبار الحرف المشكول حرفاً واحداً في الحاسوب؟

وللأجابة على ذلك فإنه يتبادر للذهن أن المواقع المخصصة في البايت الواحد البالغة 256 موقعاً لا تكفي لتمثيل الحرف مع تشكيله. ولكن ألم تقم الشركات العالمية باستحداث الرمز الدولي الموحد Unicode لتمثيل الحروف لكل اللغات بحيث يمثل الحرف الواحد ستة عشرة رقما ثنائيا (16 bits) فلماذا لا يمثل الحرف العربي لغرض التشكيل بستة عشر رقم ثنائي، وعند ذلك يمكن أن يضم الرمز ما يأتي:

أ) الحروف العربية الثماني والعشرين مع أشكال الهمزة والتاء المربوطة والألف المقصورة وهمزة الوصل والألف الخنجرية وعددها لا يزيد عن 64 رمزاً أي 6 أرقام ثنائية .(6 bits)

ب) أدوات التشكيل منفردة ومجتمعة مع الشدة ويمكن أن يخصص لها أربعة أرقام ثنائية (bits 4) حيث أن عددها لا يزيد عن 16.

جـ) رمز يشير إلى الحالة اللفظية للحرف مثل (ألف واو) الجماعة.

د) رمز يشير إلى أصل الحرف أنه كان منقلباَ عن حرف آخرمثل التاء الطويلة التي أصلها تاء مربوطة أو الألف اللينة التي كان أصلها ألفا مقصورة.

هـ) طالما أنه هناك متسع فمن الممكن أنه تكون الحروف المذكورة في الفقرة (أ) بشكل واحد ويرمز إلى شكل كتابتها (خاصة الهمزة وأنواع الألف) بحيث يشار بعدد من الأرقام الثنائية (ثلاثة أو أربعة مثلاً) إلى شكل كتابتها (على كرسي الياء أو الألف أو منفردة .. الخ).

إن هذا المقترح يناسب الحالة المشكولة للغة العربية وأية معالجة صرفية أو نحوية. ويجب أن يأخذ بنظر الاعتبار المحافظة على الاتساق مع الكتابة غير المشكولة دون تعارض.

قد يبدو هذا متناقضا مع الرمز الدولي الموحد، ولكن الحل المتفق مع الرمز الدولي الموحد سيؤدي بنا إلى استعمال 48 رقم ثنائي لتمثيل الحرف المشكول بشدة مع فتحة مثلاً، أي أن الحجم سيكون ستة أضعاف الحجم الحالي للكلمة غير المشكولة. أليس ذلك يستوجب حلاً؟

هذا المقترح مطروح للمناقشة من قبل ذوي الاهتمام باللغة العربية والحاسوب لعل حلاًمناسباً يتبلور ويأخذ طريقه في الدخول إلى حواسيب القرن الحادي والعشرين والتي ستكون أكثر ذكاء من حواسيبنا اليوم.

...والله الموفق

المصادر

1 نبيل علي: "اللغة العربية والحاسوب" (دراسة بحثية) 1988 شركة العريض.

2 Ahmed I. Abu -El Haija
Recent Advances on Standarizing Computer Arabization

3 المؤتمر الوطني الأول للحاسبات الإلكترونية وتطبيقاتها في الأردن. 2-86/11/5 في عمان

4 Al-Saleh Mohammed Amin
Standards for Arabic Characters in Informatics

5 الندوة الدولية لنقل تكنلوجيا الحاسوب – نقابة المهندسين والاتحاد العالمي للمنظمات الهندسية
23-87/11/25 عمان

6 John Clews, Language Automation Worldwide
The Development of Character Set Standards 1988
Sesmme computer Projects

7 The Unicode standard (World Wide Character Encoding)
Version 1.0, Vol.1 & 2
1990, Addison - Wesley Publishing Company. Inc

8 Sedco, Sedco Arabization Guide
1993, Amman

9 Royal Scientific Society : "Study of Arabization Solutions
for the IBM Macro Computers"

10 presented to The First Kuwait Computer Conference
March 27-29/1989, Kuwait

11 Microsoft Development Technology group, "The ABC
of converting to Unicode" 1992, 1993.

12 Microsoft Development Library, Microsoft 1992, 1993.

13 عبد القادر الكاملي: "برامج التعرف على الحروف العربية لا زالت تجهل الكثير" مجلة PC-Magazine
الطبعة العربية– السنة الثانية – العدد الخامس مايس 1996

جدول (1)

اللام \ الألف	بلا	أ	إ	آ
بلا	لا	لأ	لإ	لآ
سكون	*	لأ[1]	لإ[2]	لآ
فتحة	لا	لأ[3]	لإ[4]	لآ
كسرة	*	لأ	لإ	*
ضمة	*	لأ	*	لآ
شدة فتحة	لا	لأ	لإ	*
شدة كسرة	*	لأ	لإ	*
شدة ضمة	*	لأ	لإ	*
تنوين فتح	لا	*	*	*
تنوين ضم	*	*	*	*
تنوين كسر	*	*	*	*

حيث تشير العلامة * إلى عدم وجود ذلك باللغة العربية

والأمثلة على (1) الأخرى

و (2) الإحسان

و (3) لأنتم

و (4) تكتب كما في لئن

ملحق (1)

أطقم الحروف العربية

طوقم الحرروف

منظومة مايكروسوفت

النافذة المطورة

النافذة المطورة مع التشكيل

النافذة العالمية

النافذة العالمية مع التشكيل

النافذة الطبعة الرابع

ملحق اللغة العربية الشفاف لشركة ا.ب.م

الحروف القياسية المرقمة 449+ ASMO

الحروف العربية 701 ASMO

منظومة الدوس العربي 708

منظومة الدوس العربي 709

منظومة الدوس العربي 864

منظومة الدوس العربي 611

منظومة الدوس العربي 710 مع التشكيل

منظومة الدوس العربي 710 مع اللهجات

منظومة الدوس العربي 710 مع اللهجات والتشكيل

منظومة الدوس العربي 711

منظومة الدوس العربي 720

منظومة الدوس العربي 720 مع التشكيل

منظومة الدوس العربي 720 مع اللهجات

منظومة الدوس العربي 720 مع اللهجات والتشكيل

منظومة شركة العالمية (صخر)

منظومة المساعد العربي 786

منظومة الأمير

منظومة العربي

منظومة سيدكو (ESPRIT)

منظومة BEHIVE

منظومة EMULOG

منظومة ا.ب.م XBASIC

منظومة ARCII

منظومة CODAR-UFD

منظومة COMTERM رقم 341

منظومة COMTERM رقم 541

منظومة COMTERM رقم 348

منظومة ديتا جنرال

منظومة ا.ب.م APC
منظومة ICL-CODAR
منظومة NCR رقم 96
منظومة NCRرقم 64
منظومة HP رقم 8
منظومة HP المطورة السفلى
منظومة HP المطورة العليا
منظومة 708 مع الأطقم الفرنسية
منظومة MS WORK العربية
منظومة ARAB WORD
منظومة لوتس 1 2 3 العربية

اللغة الأردية

اللغة السندية

لغة البشتو

لغة البلوش

لغة الملايا

لغة اليوغر

لغة اللاهندا

لغة البربر